

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2002, 19:53:16 : Search time 2690 Seconds
(without alignments)
16498.803 Million cell updates/sec

Title: US-10-025-514-7

Perfect score: 1525

Sequence: 1 tctagaccatgtctggaag.....ccaactcagaagttagtcgac 1525

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pi.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pi.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	630.4	41.3	1308	6	AR111412 Sequence
2	629.4	41.3	1185	6	AR111411 Sequence
3	442.4	29.0	1345	6	AR111411 Sequence
4	438	28.7	1434	6	HUMALATB
5	436.4	28.6	1312	6	E00631
6	434.8	28.5	1584	9	BC011991
7	433.2	28.4	1352	6	AX335339 Sequence
8	433.2	28.4	1352	6	HUMALATM
9	433.2	28.4	1371	6	AX335338
10	433.2	28.4	1371	9	HSATPR1
11	433.2	28.4	1399	9	AK026174
12	433.2	28.4	1434	6	E00195
13	433.2	28.4	1434	6	E00195
14	433.2	28.4	1434	6	I04196
15	433.2	28.4	1434	6	I04272
16	433.2	28.4	1434	6	I07849
17	433.2	28.4	1435	6	AX019486
18	433.2	28.4	2478	17	AF130068
19	431.6	28.3	1337	9	HUMALAT2
20	431.6	28.3	1378	6	I02398
21	431.6	28.3	1431	9	BC015642
22	430.4	28.2	2571	17	AF113676
23	430	28.2	1185	6	AR111410
24	430	28.2	1299	6	I07949
25	430	28.2	1317	6	I00556
26	430	28.2	1378	6	I03509
27	430	28.2	1378	6	I07947
28	428.4	28.1	1308	9	BABALATA
29	428.4	28.1	1378	6	I01227
30	426.8	28.0	1185	6	A01846
31	418.4	27.4	1191	9	AB004044
32	411.2	27.0	1390	6	AX202089
33	408.8	26.8	1356	6	I36163
34	407.2	26.7	1356	6	I36161
35	404	26.5	1356	6	I36164
36	401.4	26.3	2013	6	AX472008
37	392.2	25.7	1351	10	AB000550
38	389	25.5	1372	10	AB000552
39	386.8	25.4	1351	4	SSANTIOLA
40	378	24.8	1306	10	RATATRAL
41	377.4	24.7	1352	10	S77822
42	376.4	24.7	1380	10	RATATRAL
43	376	24.7	1343	10	AB000546
44	375.6	24.6	1242	6	E13268
45	370.6	24.3	1380	4	B7ALAT

ALIGNMENTS

RESULT 1
AR111412
LOCUS
DEFINITION Sequence 6 from patent US 6127145.
ACCESSION AR111412
VERSION AR111412.1 GI:12828260
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1308)
AUTHORS Sutliff,R.D. and Rodriguez,R.L.
TITLE Production of .alpha..sub.1 -antitrypsin in plants
JOURNAL Patent: US 6127145-A 6 03-OCT-2000;
FEATURES Location/Qualifiers
1308 bp DNA linear PAT 14-FEB-2001

source	1. .1308	Score 630.4;	DB 6;	Length 1308;	
BASE COUNT	298 a	442 c	363 g	205 t	
ORIGIN	/organism="unknown"				
Query Match	41.3%;	Score 630.4;	DB 6;	Length 1308;	
Best Local Similarity	70.8%;	Pred. No. 2.1e-125;			
Matches 838;	Conservative	0;	Mismatches 346;	Indels	0;
0;					
Qy	335	GGAGACGCTCAAGGCGGCGCTCAAAAACCGACACAGTCATCAGACCAAGACCA	394		
Db	116	GAGGAGCCCGCAGGCGGACCGCCGCGGAGAGACGACACAGCAGCAGCAGCAGCA	175		
Qy	395	TCCGAGCTTTTAAATAAATTAATCAAAATTTAGCGGAATTTGCTTTTCTTTGATAGACA	454		
Db	176	CCGAGCTTCAACAAGATCACCCGGAATTTGGCGGAATTCGCTTCAGCGCTGACCGCCA	235		
Qy	455	ATTAGCTCATCAAGTAATCTACTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	514		
Db	236	GCTCGCGCACACGCTCAACTCCACACATCTTCTCAGCCCGGTGAGCATGCGCCACCGC	295		
Qy	515	TTTCGCCATGTTAGTTTATAGTCTAAAGCGATACCCATGACGAGATTTTGAAGGTTT	574		
Db	296	CTTCGCCATGCTGCTCCCTGGGTACCAAGCGGACACCCAGCAGAGATCTCGAAGGGCT	355		
Qy	575	AACTTTTAAATTTGACCGAAATCCAGAGGCCCAATTCAGAGGGTTTTCAGAGGTTGTT	634		
Db	356	GAACTTCAACCTGAGGAGATCCCGGAGGCGCAGATCCAGAGGGCTTCCAGGAGCTGCT	415		
Qy	635	GAGAACTTTGAATCAACCTGATTTCTCAATTTCAATTAATCACTACTCGTAACGGTTTATTTT	694		
Db	416	CAGGAGCTCAACGACCGGAGCTCCAGCTCCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTC	475		
Qy	695	GTCGAGGTTTAAATTTGGTTGACAAATTCCTAGAGACGCTCAAGAACTATATATAG	754		
Db	476	GTCGAGGCTCAAGCTGCTGTAAGTTCTCTGGAGGACGCTGAAGAGCTCTACCACTC	535		
Qy	755	TGAGGCTTTTACCGTTAAATTTTGGTGATAGTGAAGGAGCTAAAAGCAATTAATGATTA	814		
Db	536	CGAGGCTTCAACGCTCACTTCGGGACACCGAGGAGGCGCAAGAGCAAGATCAACGACTA	595		
Qy	815	TGTTGAGAAAGGACCCAGGCTAAGATCGTTGACCTAGTTAAAGATTTAGATCGTGATAC	874		
Db	596	CGTCGAGAGGGGACCCAGGCGAAGATCGTGACCTGCTCAAGGAATTTGGACAGGACAC	655		
Qy	875	CGTCTTCGACTAGTTAACTATATTTTTCAGAGGTAAGTGGGAACTCTCTTCGAGGT	934		
Db	656	CGTCTTCGCTCGTCAACTACATCTTCTTCAAGGCAAGTGGGAGCGCCGCTTCGAGGT	715		
Qy	935	TAAAGATCTGAAGAGAAAGATTTTCATGTTGATCAAGTTACTACTGTCAGAGTTCCAT	994		
Db	716	GAAGGACACCGGAGGAGGACTTCCAGCTGACACGAGTCAACCGCTCAAGGTCCTCGAT	775		
Qy	995	GATGAAAGACTGGGTATGTTCAATTTCAACATTTCAACATTTCAAGTTCTTGGGCTCT	1054		
Db	776	GATGAGAGGCTCGGATGTTTCAACATTTCAACATTTCAACATTTCAAGTTCTTGGGCTCT	835		
Qy	1055	ATTAAATGAAGTATTTAGGTAACTGCTACTGCTATTTTTCAGAGGAAAGGTAAGCT	1114		
Db	836	CCTCATGAAGTACTTGGGAAACCCAGCGCATCTTCTTCTCGCGGAGGAGGCAAGCT	895		
Qy	1115	TCAACATTTAGAGTAAGTTGACTCATGACATTTACTAATTTTATAGAACGAGGA	1174		
Db	896	CCAGCAGCTGGAGAACGAGCTGACGACGACATCATCAGAGTTCTCTCGGAGAACGAGA	955		
Qy	1175	TCGTCGTAGCGCTCTCTGCACTGCGCAAGTTTAAAGTATCACCGGTACTTACGACTTAA	1234		
Db	956	CAGGCGCTCCGCTAGCTCCACCTCCGAGAGCTGAGCATCACCGGACGCTACGACCTGAA	1015		
Qy	1235	ATCTGTTTTAGCGGAGTTAGTATTAACAAAGTTTTTTTCAAGCGTCCGATTTAGTGG	1294		
Db	1016	GAGCGTGTGGGCCAGCTGGGCATCAGAAAGTCTTCAGCAACGCGGCGGAGCTCTCCCG	1075		

RESULT 2
 AR111411
 LOCUS
 DEFINITION Sequence 3 from patent US 6127145.
 ACCESSION AR111411
 VERSION AR111411.1 GI:12828259
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 1185)
 AUTHORS Sutilff,T.D. and Rodriguez,R.L.
 TITLE Production of .alpha. .sub.1 -antitrypsin in plants
 JOURNAL Patent: US 6127145-A 3 03-OCT-2000;
 FEATURES
 source
 1..1185
 /organism="unknown"
 BASE COUNT 276 a 396 c 334 g 179 t
 ORIGIN
 Query Match 41.3%; Score 629.4; DB 6; Length 1185;
 Best Local Similarity 70.8%; Pred. No. 3.4e-125;
 Matches 837; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

Qy 1295 TGTTACTGTAAGAAGCTCCATTAATAATTGAGTAAGCTGTTTCAAAAGCCGCTTAACTAT 1354
 Db 1076 CGTGACGGAGGAGGCGCCCTGAAGCTCTCAAGCCGCTGCAAAAGCGGCTGCTCAGCAT 1135
 Qy 1355 TGATGAAAAGGTTACCGAGGCGCGCGCTATGTTCTTGGAAAGCTATTTCCTCAATGAGCAT 1414
 Db 1136 CGACGAGAAGGAGGAGGAGCTCCCGGGCCATGTTCTTGGAGGCCATCCCATGTCCAT 1195
 Qy 1415 TCCACCAAGAAGTTAAATTTAATAAACCATTCGTTTTCGTATGATCGAGCAGAACACTAA 1474
 Db 1196 CCCGCCGAGGTCAAGTTCAACAAGCCCTTCGTTCTTCTGATGATCGAGCAGAACAGAA 1255
 Qy 1475 AAGCCCATGTTTATGAGGTAAAGTTGTCAACCACTCAGAAGT 1518
 Db 1256 GAGCCCTCTTTCATGGGGAAGGTCTCAACCCACGACGAGT 1299

Qy 336 GAAGACCTTCAAGGCGGAGCGCTCAAAAACCGACACGATCATCAGACCAAGACCAT 395
 Db 1 GAGGACCCGAGGCGGAGCGCCCGCAGAAAGCCGACACCGCCACGACGACGACGAC 60
 Qy 396 CGACCTTTTAAATAAATTAATCTCAAAATTTAGCCGAATTTGCTTTTCTTTGTATAGCAA 455
 Db 61 CCAGCTTTCAACAAGATCACCCGGAATTTGGCCGAATTCGCTTCAGCTTCAGCGCAG 120
 Qy 456 TTAGCTCATCAAGAGTAATTTCTACTAACAATTTTTCCTAGTCTGTTTCTTATTCGCACTGCT 515
 Db 121 CTCGCGACACGATCCCAACTCCCAACATCTTCTTCAGCCGGTGAGCATCGCCACGCC 180
 Qy 516 TTCGCATGTTGAGTTTAGTACTAAGCCGATACCCATGACGAGATTTTATAGAGGTTTA 575
 Db 181 TTCGCATGCTGTCCTGGGTACCAAGCGGAGACCCAGCAGAGATCTCGAAGGCTG 240
 Qy 576 AACTTTAATTTGACGAAATCCCGAAGCCCAAAATTCAGAGGGTTTTCAGAGTTGTTG 635
 Db 241 AACTTCAACTGACGAGATCCCGGAGGCGCAGATCCAGGAGGCTTCAGAGGCTGCTC 300
 Qy 636 AGAATTTGAATCAACCTGATCTCAATTTCAATTTCACTTGGTAACGGTTTATTTTGT 695
 Db 301 AGGACGCTCAACGAGCGGACTCCCGAGCTCCAGCTCACCACCGGACGCGGCTCTCTCG 360
 Qy 696 TCTGAAGGTTTAAATTTGGTTGACAAATTCCTAGAGAGCTCAAGAAACTATATCATAGT 755
 Db 361 TCCGAGGCGCTCAAGCTCGTCGATAGTTCTTGGAGAGCTGAAGAAGCTCTTACCACTC 420
 Qy 756 GAGGCTTTTACCGTTAAATTTTGGTGATGATGAGAGAGCTAAAAGCAATTAATGATTA 815
 Db 421 GAGCGTTTACCGCTCAACTTCGGGGACACCGAGGAGGCGCAAGAGCAGATCAACGACTAC 480


```
Db 452 CCTGAAGCTAGTGGATAAGTTTGGAGGATGTTAAAGTGTACCACCTCAGAAGCCTT 511
Qy 764 TACCCTTAATTTTGTGATGACTGAGGAGCTTAAAGCAAAATTAATGATGATTTGAGAA 823
Db 512 CACTGTCACTTCTGGGGATCAGAGAGGCGCAAGAACAGATCAACGATTAACGTTGGAGAA 571
Qy 824 AGGCCACCCAGGTAAGATCGTTGACCTAGTTAAAGAAATTAGATCGTGATACCGCTTCGC 883
Db 572 GGGTACTCAAGGAAATTTGTTGATTTGGTCAAGGAGCTTCACAGAGACACAGTTTTCG 631
Qy 884 ACTAGTTAACTAATTTTCAAGGGTAAGTGGGAAGCTCTTCGAGGTTAAAGATAC 943
Db 632 TCTGTGTAATACATCTCTTTAAAGGCAATGGGAGAGACCTTTTGAAGTCAAGGACAC 691
Qy 944 TCAAGAGGAAGATTTTCATGTTGATCAAGTTTACTACTCTCAAGTTCCAAATGATGAAAG 1003
Db 692 CGAGACAGGAGCTTCCAGCTGGACAGCTGACCCAGTGAAGTCCCTATGATGAAGCG 751
Qy 1004 ACTGGGTATGTTCAATATTCACATTCACAAATTAAGTTCTTGGGTCCTTAAATGAA 1063
Db 752 TTTAGGCACTGTTTAAACATCCAGCACTGTAAGAAAGCTGTCAGCTGGGTACTGCTAAATGAA 811
Qy 1064 GTATTAGTAACTGCTACTGCTATTTTCTTTTACACAGCAAGTAACTTCAACATTT 1123
Db 812 ATACCTGGCAATGGCACCGCATCTTCTCTACCTGATGAGGGGAACCTACAGCACCT 871
Qy 1124 AGAAGATGAGTTGACTCATGACATGATTAATTAATTTTATAGAACAGGATCGTCGTAG 1183
Db 872 GGAATATGAACTCACCCAGGATATCATCACCAGTTCTCGGAATGAAGACAGAGTGC 931
Qy 1184 CGCTCTCTGACCTGCCAAAGTTAAGTATCACCGGTACTTACGACTTAAATCTGTTTT 1243
Db 932 TGCCAGCTTACATTTACCCAACTGCTCATTACTGGAACCTATGATCTGAAGAGCGTCT 991
Qy 1244 AGCCAGTTAGTATTAACAAAGTTTCTTAAACGGTCCGATTTGAGTGTGTTACTGA 1303
Db 992 GGGTCAACTGGGCATCACTAAGGCTCTCAGCAATGGGCTGACCTCTCCGGGGTCACAGA 1051
Qy 1304 AGAAGCTCCATTAATTAAGTAAAGCTGTTTCAAAAGCGCTTAACTATTGATGAAA 1363
Db 1052 GGAGCACCCCTGAAGCTCTCAAGCGCTGCATGAAGCTGTGTCACCATCGAGAGAA 1111
Qy 1364 GGGTACCGAGCGCGCGGCTGATGTTCTCGGAAGCTATCAATGAGAGATTCACACAGA 1423
Db 1112 GGGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCAATGTCTATCCGCCCA 1171
Qy 1424 AGTTAAATTAATAACCATTCGTTTTTCTGTGATGATCGAGCAGACACTAAAGCCCAT 1483
Db 1172 GGTCAAGTTCAACAAACCCCTTGTCTCTTAATGATGAACAAAATACCAAGTCTCCCT 1231
Qy 1484 GTTTATGGGTAAGGTTGTCAACCCCAACTCAGAAAGTA 1519
Db 1232 CTTTCATGGGAAAGTGGTGAATCCCAACCCAAAATA 1267

RESULT 4
E00631
LOCUS E00631 1434 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA encoding human alpha-1 antitrypsin mutain.
ACCESSION E00631
VERSION E00631.1 GI:2168910
KEYWORDS E00631.1 JP 1986012289-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1434)
AUTHORS Maagetusto,W.I. and Guren,H.K.
TITLE REGION SPECIFIC MUTATION INDUCTION IN ALPHA-1-ANTI-TRYPSIN
JOURNAL Patent: JP 1986012289-A 1 20-JAN-1986;
CHIMO JIENETEITSUKUSU INC
COMMENT OS human
```

```
PN JP 1986012289-A/1
PD 20-JAN-1986
PF 14-MAR-1984 JP 1985051553
PR 14-MAR-1984 US 84 589410, 07-MAR-1985 US 85 709382 PI
MAAGARETSUTO WAI INZURIL, GUREN HITOSHI KAWASAKI PC
C12N15/00,A61K35/74,A61K37/64,C07H21/04,C07K15/04, PC
C12P21/02//A61K37/04,
PC C12N9/99,(C12P21/02,C12R1:865);
CC strandedness: Both;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue_type=livr;
FH Key Location/Qualifiers
FT CDS 47..1303
FT /product='alpha-1 antitrypsin' FT variation
FT replace(1190..1193,'GGG')
FT /note='amino acid 358=Ala'
FT variation
FT replace(1190..1193,'GTG')
FT /note='amino acid 358=Val'
FT variation
FT replace(1190..1193,'GGG')
FT /note='amino acid 358=Gly'
FT variation
FT replace(1190..1193,'TTC')
FT /note='amino acid 358=Phe'
FT variation
FT replace(1190..1193,'AGA')
FT /note='amino acid 358=Arg'
FT variation
FT replace(1190..1193,'AGG')
FT /note='amino acid 358=Arg'
FT variation
FT replace(1190..1193,'AAG')
FT /note='amino acid 358=Lys'
FT variation
FT replace(1142..1144,'AAG')
FT /note='amino acid 342=Lys'.

FEATURES
source
1..1434
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 369 a 426 c 346 g 293 t
ORIGIN
Query Match 28.7%; Score 438; DB 6; Length 1434;
Best Local Similarity 59.9%; Pred. No. 5,1e-84;
Matches 732; Conservative 0; Mismatches 490; Indels 0; Gaps 0;
Qy 298 TGTGTGTAAGTCTGTGTTTCCCGAGTCAAGCCATGGAAGACCCCTCAAGGCGAGCGCG 357
Db 81 TGGCAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 140
Qy 358 CTCAAAAACCGCACACAGTCATCCAGCAACCAAGACCATCCGACTTTTAAATAAAATTAATC 417
Db 141 CCCAGAAGACAGATACATCCACCATGATCAGGATCACCCACCTTCAACAAATCACC 200
Qy 418 CAAATTTAGCGGAATTTGCTTTTCTTTTATAGACAATTTAGCTCATCAAGTAATTTA 477
Db 201 CGAACCTGGCTGAGTTCGCTTCAGCCTATACGCCAGTATACGCCAGTGGCACACAGCA 260
Qy 478 CTAACATTTTGTAGTCTGTTTCTATTCGCACTGCTTTCGCACTGCTGAGTTAGTTAG 537
Db 261 CCAATATCTTCTTCTCCCGAGTGAGCATCGCTACAGCTTTTGAATGCTCTCCCTGGGA 320
Qy 538 CTAAGCGGATACCCATGACGAGATTTTGAAGGTTTAAAGCTTTTAAATTTGACGAAATCC 597
Db 321 CCAAGGCTGACATCAGATGAATCTCGAGGCGCTGAATTTCAACCTCAGGAGATTC 380
Qy 598 CAGAAGCCCAAAATTCAGGAGGTTTTCAGAGTTTGTGAGAATTTTGAAGTTCATCAACCTGATT 657
Db 381 CGGAGGCTCAGATCCATGAAGGCTTCCAGGAAGTGTCTTACCTTACCTCAACAGCAGCA 440
Qy 658 CTCATTTGCAATTAACCTACTGTTAAACGTTTATTTTGTCTGAAGTTTAAATTTGTTG 717
Db 441 GCCAGCTCCAGCTGACCCAGCGCAATGGCTGTCTTCTGAGCGAGGCGCTGAAGCTAGTG 500
Qy 718 ACAATTTCCITAGAGAGCTCAAGAAACTATATCATAGTGAGGCTTTTACCGTTAAATTTG 777
```

```

Db 501 ATAAGTTTGGAGATGTTAAAGAGTTGTACCACTCAGAAGCCTTCACTGTCAACTTCG 560
Qy 778 GTGATACAGAGAGCTAAAAGCAAAATTAATGATTAATGTTGAGAAGGCCACCCAGGGTA 837
Db 561 GGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAAGTGTGAGAGAGGGTACTCAAGGGA 620
Qy 838 AGATCGTTGACCTAGTTAAAGAAATAGATCGTGATACCGTCTTCGCACTAGTTAACTATA 897
Db 621 AAATTTGGATTTGGTCAAGGAGCTTGACAGAGACACAGTTTTTGTCTGTGTGAATTACA 680
Qy 898 TTTTTTCAAGGGTAAGTGGGAAGCTCTCTTCGAGGTTAAAGATACATGAAGAGGAAGATT 957
Db 681 TCTTCTTTAAAGCAATGGGAGAGACCTTTTGAAGTCAAGGACACCGAGAAGAGACT 740
Qy 958 TTCATGTTGATCAAGTTACTACTGTCAAAGTTCCAAAGTTGAAAGAGACTGGGTATGTTCA 1017
Db 741 TCCACGTTGGACCAAGTGCACCACTGAGGTCGCTGATGAATATACCTGGGCAATG 800
Qy 1018 ATATTCAACATTTGCAAAAATTAAGTTCTTGGGTCCTTATTAAAGAGATTTAGGTAAAG 1077
Db 801 ACATCCAGCACTGTAAAGAGCTGTGAGCTGGGTCCTGCTGATGAATATACCTGGGCAATG 860
Qy 1078 CTACTGCTATTTTTTTTTTACCAGACGAAGGTAAAGCTTCAACATTTAGAAATGAAGTTGA 1137
Db 861 CCACCGGCATCTTCTCTGCTGATCAGGGGAAACTACAGCACTTGAAATGAAGTTCA 920
Qy 1138 CTCATGACATTAATTAATTTTACAGAACGAGGATCGTCTGAGCGCTTCTCTGCAAC 1197
Db 921 CCCAGGATATCATCACCAGTTCTCTGGAATTAAGAGGAGGAGGCTCTGCGAGCTTACATT 980
Qy 1198 TGCCAAAGTTAAGTATCACCGGCTACTTACGACTTAAATCTGTTTAAAGGCGAGTTAGGTA 1257
Db 981 TACCAAACTGTCCATTACTTGAACCTATGATCTGAAGAGCGTCTGGGTCAACTGGGCA 1040
Qy 1258 TTACCAAGTTTCTTAAAGCGTCCGATTTGAGTGGTGTACTGAGAAAGCTTCCATTAA 1317
Db 1041 TCACAAAGGCTTTCAGCAATGGGGCTGACCTCTCCGGGGTCACAGAGAGGACACCCCTGA 1100
Qy 1318 AATTGAGTAAAGCTGTTTACAAAAGCGCTCTTAACATTTATGATGAAAGGGTACCGAGGCCG 1377
Db 1101 AGCTCTCAAGGCGTGCATAAGCGTGTGCTGACCATCGACGAGAAAGGACTGGAAGCTG 1160
Qy 1378 CCGGCGTATGTTCTCGGAAGCTATTCCAATGAGCATTCACACAGAGTTCATTAATTAATA 1437
Db 1161 CTGGGCCATGTTTTTAGAGGCCATACCCATGTCTATCCCGCCGAGGTCAGTTCAACA 1220
Qy 1438 AACCATTCGTTTTCTGATGATCGACGACAGCACTAAAGCCCATTTGTTATGGGTAAAG 1497
Db 1221 AACCTTTGCTCTTAAATGATTGAACAAATACCAAGTCTCCCGCTGTTTCATGGGAAAAG 1280
Qy 1498 TTGTCAACCACTCAGAAGTA 1519
Db 1281 TGGTGAATCCCAACCAAAAATA 1302

RESULT 5
102706
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
```

```

source 1..1312
BASE COUNT 339 a 368 c 324 g 281 t
ORIGIN
Query Match 28.6%; Score 436.4; DB 6; Length 1312;
Best Local Similarity 59.8%; Pred. No. 1.1e-83;
Matches 731; Conservative 0; Mismatches 491; Indels 0; Gaps 0;
Qy 298 TGTGTGGTAGTCTCTGTGTTTCCCCAGTCAAGGCCATGGAAGACCTCAAGGGAGCGCG 357
Db 62 TGCAGGCGTGTCTGCTGCCCTGGTCCCTGTCTCCCTGGCTGAGGATCCCGAGGAGATGCTG 121
Qy 358 CTCAAAAAACCCAGACAGCAGTCATCAGCAGCAAGACCATCCGACTTTTAAATAAAATFAC 417
Db 122 CCCAGAAGACAGATACATCCACCATGATCAGGATCACCACCACTTCACAAGATCACCC 181
Qy 418 CAATTTAGCCGAAATTTGCTTTTCTTTGTATAGACAAATTAGCTCATCAAAAGTAATCTTA 477
Db 182 CCAACTTGGCTGAGTTTCGCCCTTCAGCCTATACCGCCAGCTGGCACACCAAGTCCAACAGCA 241
Qy 478 CTAACATTTTTTTAGTCTCTGTTTCTATTGCCACTGCTTTCGCCATGTTGAGTTTAGGTA 537
Db 242 CCAATATCTCTTCTCCCGAGTGAGGATCGCTACAGCCTTTCGAATGCTCTCCCTGGGGA 301
Qy 538 CTAAGCCGATACCCATGACGAGATTTTAAAGAGTTTAAACTTTTAAATTTGACCGAAATCC 597
Db 302 CCAAGGCTGACACTCAGCATGAAATCCTGGAGGGCTGAATTTCAACCTCACGGAGATTC 361
Qy 598 CAGAGCCCAATTCACGAGGGTTTTTCAAGAGTTTGTGAGAACTTTGAATCAACCTGATT 657
Db 362 CGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTCCGTACCTCAACCGACGACACA 421
Qy 658 CTCAATTTGCAATTAACACTACTGTGTAACGGTTTTATTTTGTCTGAAGTTTAAATTTGGTTG 717
Db 422 GCCAGCTCCAGCTGACCCCGCAATGGCTGTCTCCTCAGCGAGGGCTGAAGCTAGTGG 481
Qy 718 ACAAAATTCCTAGAACAGCTCAAGAACTATATCATAGTAGAGCTTTTACCGTTAAATTTTG 777
Db 482 ATAAGTTTTTGGAGGATGTTTAAAGAGTTGTACCACTCAAGAGGCTTCACTGTCAACTTCG 541
Qy 778 GTGATCTAGGAGAGCTAAAAAGCAAAATTAATGATTATTTGTGAGAAAGGCCACCGAGGTA 837
Db 542 GGGACACCGAAGAGGCCAAGAAACAGATCAAGATTACGTGGAGAGGGTACTCAAGGGA 601
Qy 838 AGATCGTTGACCTAGTTTAAAGAAATTAGATCTGTGATACCGTCTTCGCACTAGTTTAACTATA 897
Db 602 AAATTTGTGGATTTGGTCAAGGAGCTTGACAGAGACACAGTTTTTGTCTGTGTAATTAACA 661
Qy 898 TTTTTTTTCAAGGGTAAGTGGGAACGCTCTTCGAGGTTAAAGATACCTCAAGAGGAGGATT 957
Db 662 TCTCTTTAAAGGCAATGGGAGAGACCTTTGAAGTCAAGGACACCGAGGAGGAGGACT 721
Qy 958 TTCATGTTGATCAAGTTACTACTGTCAAAAGTTCCAATGATGAAAAGCTGGGTATGTTCA 1017
Db 722 TCCACGTGGACCAAGTGACCCCGTGAAGGTGCTCTATGATCAAGCGTTTAGCATGTTTA 781
Qy 1018 ATATTCAACATTTGCAAAAATTAAGTTCTTGGGTCTTATTAAATGAAGTATTTAGGTAAAG 1077
Db 782 ACATCCAGCATTTGAAGAAAGCTGTCAGCTGGGTGCTGCTGATGAAATFACCTGGGCAATG 841
Qy 1078 CTACTGCTATTTTTTTTTTACAGAGGAGTAAAGCTTCAACATTTTAGAGAAATCAGTTGA 1137
Db 842 CCACCGCCATCTTCTTCTGCTGATGAGGGGAACTACAGACCTCGGAAATGAAGTCA 901
Qy 1138 CTCATGACATTTACTATAAATTTTATAGAGAACGAGGATCGTCGTAGGCGCTTCTCTGCACC 1197
Db 902 CCCAGGATATCATCACCAGTTTCTCTGGAAAATGAAGACAGAAAGCTCTGCCAGCTTACATT 961
Qy 1198 TGCCAAAGTTTAAAGTATCACCGGTACTTACGACTTAAATCTGTTTAAAGGCGAGTTAGGTA 1257
Db 962 TACCAAACTGTCCATTACTGGAACCTATGATCTGAAGAGGCTCTAGGTCAACTGGGCA 1021
```


QY	898	TTTTTTTCAAGGTAAGTGGGACGCTCTTTTCGAGGTTAAAGATACTGAAGAGGAGATT	957
Db	654	TTTTTTTAAAGCAATGGGAGAGACCTTTTGAAGTCAAGGACACCGAGGAGGACT	713
QY	958	TTTCATGTTGATCAAGTTACTACTGTCMAAGTTTCCAAATGATGAAGAAAGACTGGGTATGTTCA	1017
Db	714	TTCCACGTGGACGAGGTGACACCGTGAAGGTGCTATGATGAAGGTTTAGGCATGTTTA	773
QY	1018	ATATTCAACATTTGCAAAAATTAAGTTCTTTGGTCTTATTAATGAAGTATTAGGTAAGC	1077
Db	774	ACATCCAGCAGTGAAGAAGCTGTCCAGCTGGGTGCTGCTGATGAATACTCTGGGCAATG	833
QY	1078	CTACTGCTATTTTTTTTACCAGAGGAGTAAAGTAAAGCTTCAACATTTAGAGAATGAGTTGA	1137
Db	834	CCACCGCCATCTCTCTCCCTGCTGATGAGGGAAACTACAGCACCTGGAAATGAATCA	893
QY	1138	CTCATGACATTTACTTAATTTTATAGAGACGAGATGCTGAGCGTTCTCTGCAAC	1197
Db	894	CCACGATATCATCACCAAGTTTCCGGGAAATGAAGACAGAAAGGCTGCCAGCTTACAT	953
QY	1198	TGCCAAGTTAAGTATCACCGGTACTTACGACTTAAATCTGTTTTAGGCCAGTTAGGTA	1257
Db	954	TACCCAACTGCTTACTTACTGGAACCTATGATCTGAAGAGCTCTGGTCAACTGGGCA	1013
QY	1258	TTACCAAGTTTCTTAACGGTGCGGATTTGAGTGGTTTACTGAAGAAGCTCCATTAA	1317
Db	1014	TCACTAAGGCTTTCAGCAATGGGCTGACCTCTCCGGGTACACAGGAGGACCCCTGA	1073
QY	1318	AATTGATGAAGCTGTTCAAGAAGCGCTTAACTATTGATGAAGAAAGGTTACCGAGCGC	1377
Db	1074	AGCTCTCAAGCGCGTGCATAAGGCTGTGCTGACCATCGACGAGAAAGGACTGAAGCTG	1133
QY	1378	CGGGGCTATGTTCTCGGAAGCTATTCGAATGAGCATTCACCAAGATTCACCAAGTTAAATTAATA	1437
Db	1134	CTGGGCGCATGTTTTAGAGGCCATACCCATGCTATCCCCCGGAGGTCAAGTTCAACA	1193
QY	1438	AACCATGCTTTTCTGATGATCGAGCAGACACTAAAGAGCCATTTGTTATGGGTAAGG	1497
Db	1194	AACCCCTTTGCTCTTAAATGATGAACAAATATCCAAAGTCTCCCTCTCTATGGGAAAG	1253
QY	1498	TTGTCAACCAACTCAGAGTA	1519
Db	1254	TGGTGAATCCCAACCAAAATA	1275
RESULT 9			
AX335338			
LOCUS	AX335338	1371 bp	DNA
DEFINITION	Sequence 5847 from Patent WO0194629.	linear	PAT 09-JAN-2002
ACCESSION	AX335338		
VERSION	AX335338.1	GI:18126057	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,		
JOURNAL	Horrikan, S., Soppet, D.R. and Weaver, Z.		
FEATURES	Cancer gene determination and therapeutic screening using signature		
source	gene sets		
BASE COUNT	Patent: WO 0194629-A 5847 13-DEC-2001;		
ORIGIN	Avalon Pharmaceuticals (US)		
	Location/Qualifiers		
	1..1371		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	350 a 388 c 340 g 293 t		
Query Match	28.4%;	Score 433.2;	DB 6;
Best Local Similarity	59.7%;	Pred. No. 5.5e-83;	Length 1371;
Matches 729;	Conservative 0;	Mismatches 493;	Indels 0;
	Gaps 0;		

QY	298	TGTGTGTAAGTCTCTGTGTTTCCCCAGTCAAGGCCATGGAAGACCCCTCAAGGGGACGCCG	357
Db	73	TGGCAGGCGCTGTGTGCTGCTCCCTGCTGAGGATCCCGAGGAGATGCTG	132
QY	358	CTCAAAAACCGACACAGTCTATCAGCAGCAAGACCATCCGACCTTTTAAATAAATTAATC	417
Db	133	CCGAGAAGACAGATACATCCACCATGATCAGGATCACCCACCTTCAACAAGATCAACC	192
QY	418	CAAAATTTAGCCGAATTTGCTTTTCTTTGTATAGACAATTAGCTCATCAAGTAATTTCTA	477
Db	193	CAAACCTGGCTGAGTTCGCTTACGCCCTTACGCCAGCTGAGGACACAGCTCCCAACAGCA	252
QY	478	CTAACATTTTTTTTAGTCCCTGTTTCTATGCCACTGCTTTTCCGCTATGTTAGTTAGGTA	537
Db	253	CCAATACTCTCTTCCCCAGTGAGCATCGCTACAGCTTTTGCATGCTCTCCCTGGGGA	312
QY	538	CTAAGCGGATPACCATGACGAGATTTTGAAGGTTTAAACTTTTAACTTTGACCCGAAATCC	597
Db	313	CCAAGGCTGACACTCAGATGAATCCTGGAGGCTGAAATTTCAACCTCAGGAGATTC	372
QY	598	CAGAAGCCCAAAATTCACGAGGTTTTCAAGAGTTGTTGAGAATTTGAATCAACCTGATT	657
Db	373	CGGAGGCTCAGATCCATGAAGGCTTCCAGGAACCTCCCTCCGTACCTCAACAGCCAGCA	432
QY	658	CTCAATTTGCAATTAACCTACTGCTAACGTTTATTTTGTCTGAGGTTTAAATTTGGTTG	717
Db	433	GCCAGCTCCAGCTGACCCAGGCAATGGCTGTTTCTCAGCGAGGGCTCAAGCTAGTGG	492
QY	718	ACAAATTTCTAGAGAGCTCAAGAAACTATATCATAGTGAAGGCTTTTACCGTTAAATTTG	777
Db	493	ATAAGTTTTTGGAGGATGTTTAAAGTTGTACCACTCAGAAGCTTCACTGTCACTTCG	552
QY	778	GTGATCTAGGAAAGCTTAAAGCAAAATTAATGATATGTTGAGAAAGGACCCAGGTA	837
Db	553	GGGACACCGAAGAGGCCAAGAAACAGATCAACGATTAACGTGAGAGAGGTTACTCAAGG	612
QY	838	AGATCGTTGACCTAGTTTAAAGAAATAGATCGTATACCGTCTTCCGCTAGTTAACTATA	897
Db	613	AAATTTGGATTTGGTCAAGGAGCTTGACAGAGACACAGTTTGTCTCTGGTGAATTTACA	672
QY	898	TTTTTTTCAAGGTAAGTGGGAAAGCTCTTTTCGAGGTTTAAAGATACTGAAGAGAAAGATT	957
Db	673	TCCTTTTAAAGCAATGGAGAGACCTTTTGAAGTCAAGGACACCGAGGAGAGGACT	732
QY	958	TTCAATGTTGATCAAGTTACTACTGTCAAAGTTCATGATGAAGAAAGACTGGGTATGTTCA	1017
Db	733	TCCACGTGGACAGGTGACCCAGCTGAAAGTGCCTATGATGAAGCGTTTAGGCATGTTTA	792
QY	1018	ATATTCAACATTCGAAAAAATTAAGTTCTTTGGGTCTTATTAAATGAAGTATTTAGGTAACG	1077
Db	793	ACATCCAGCACTGTAGAAGCTGTCCAGCTGGGTGCTGCTGATGAAGTAACTCTGGGCAATG	852
QY	1078	CTACTGCTATTTTTTTTTTACCAGAGAAAGTAAAGCTTCAACATTTAGAGAAATGAGTTGA	1137
Db	853	CCACGCCATCTCTCTCTCCCTGATGAGGGGAAACTACAGCACTTGGAAAAATGAAGTCA	912
QY	1138	CTCATGACATTTTACTAAATTTTGTAGAACAGAGATTCGCTAGCGCTTCTCTGCACC	1197
Db	913	CCACGATATCATCAACCAAGTTCTCTGGAAATGAAGACAGAGGTTCTGCGAGCTTACAT	972
QY	1198	TGCCAAAAGTTAAGTATACCGGTTACTTACGACTTAAATCTGTTTGGCCAGTTAGGTA	1257
Db	973	TACCAAACTGCTCCATTTACTGGAACCTATGATCTGAAGAGGCTCTGGGTCAACTGGGCA	1032
QY	1258	TTACCAAGTTTTTTTCTAACGGTGCCGCTTTGAGTGGTGTGTTACTGAAGAGCTTCCATTAA	1317
Db	1033	TCATAAGGCTTTTACGCAATGGGGCTGACCTCTCCGGGGTACAGAGAGGACCCCTGA	1092
QY	1318	AATTGATGAAGCTGTTCAAAAAGCCGCTTAACTATTGATGAAGAAAGGTTACCGAGCGC	1377
Db	1093	AGCTCTCCAAAGGCGGTCATGAAGGCTGTGCTGACCATCGACGAGAAAGGACTGAAGCTG	1152


```

Db 973 TACCCAAACTGTCTACTTGAACCTATGATCTGAAGACGCTCTGGGTCAACTGGGCA 1032
QY 1258 TTACCAAAGTGTCTTACAGGTGCGGATTTGAGTGGTGTACTGAAGAGCTCCATTAA 1317
Db 1033 TCACTAAGGTCTTCACAGATGGGGTGACCTCTCCGGGGTCACAGAGGAGGCCCTGA 1092
QY 1318 AATTGAGTAAAGTGTTCACAAAGCGCTCTTAACCTATTGATGAAAAGGTACCGAGGCG 1377
Db 1093 AGCTCTCCAAAGCGGTGCATAAGGTGTGCTGACCATCGACGAGAAAGGAGCTGAAGCTG 1152
QY 1378 CGGGCGCTATGTTCTGGAAGCTATTCCAAATGAGCATTCACACAGAAGTTAAATTAATA 1437
Db 1153 CTGGGCCCATGTTTTAGAGGCGCATACCCATGTCTATCCCCCGAGGTCAAGTTCAACA 1212
QY 1438 AACCTATGCTTTTCTGATGATCGAGCAGACACACTAAAGCCCATGTTTATGGGTAAGG 1497
Db 1213 ACCCTTTGCTCTCTTAATGATGTAACAAATACCAAGTCTCCCTCTTCATGGGANAAG 1272
QY 1498 TTGTCAACCCAACTCAGAAAGTA 1519
Db 1273 TGGTGAATCCCAACCAATA 1294

```

RESULT 11

AK026174
 LOCUS
 DEFINITION Homo sapiens cDNA: FLJ2521 fls, clone HRC12386, highly similar to AF113676 Homo sapiens clone FUB2803 PRO0684 mRNA.

ACCESSION AK026174

VERSION oligo capping: fls (full insert sequence).

KEYWORDS Homo sapiens primary human renal epithelial cells cDNA to mRNA,

SOURCE clone_lib:HRC clone:HRC12386.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (sites)
 Kawabata,A., Hiki,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project

TITLE

Unpublished

2 (bases 1 to 1399)

Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.
 Direct Submission
 Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdnal@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- and 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

source

1. .1399

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="HRC12386"

/cell_type="primary human renal epithelial cells"

/clone_lib="HRC"

/note="cloning vector pME18SFL3"

1. .1399

/note="highly similar to AF113676 Homo sapiens clone

FLB2803 PRO0684 mRNA"

BASE COUNT

378 a 395 c 329 g 297 t

ORIGIN

```

Query Match      28.4%   Score 433.2; DB 9; Length 1399;
Best Local Similarity 59.7%   Pred. No. 5.5e-83;
Matches 729; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

QY 298 TGTGTGTGAAGTCTCTGTGTTTCCCGAGTCAAGGGCCATTGGAAGACCTCAAGGCGACGCCG 357
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 78  TGGCAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 137
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 358 CTCAAAAACCGACGACGATCATCAGACCAAGACCATCCGACGCTTTTAATAAAATTAATC 417
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 138 CCCAGAGACAGATACATCCACCATGATCAGGATCAACCAACCTTCAACAAGATCAACC 197
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 418 CAATTTAGCCGAATTTGCTTTTCTTTGATATAGACAATTAGCTCATCAAGATTAATCTA 477
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 198 CCAACTGGCTGAGTTGCGCTTACGCTTACTGCGAGCTGGCACACGATCCCAACAGCA 257
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 478 CTAACATTTTTTTTAGTCTCTGTTTCTATTGCCACTGCTTTTCGCCCATGTTTGAGTTTAGGTA 537
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 258 CCAATATCTTCTTCCCGAGTACGCTACAGCCTTTTGCAATGCTCTCTCCCTGGGA 317
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 538 CTAAGCGGATACCCATGACGAGATTTAGAGGTTTAAACTTTAAATTTGACCGAAATCC 597
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 318 CCAAGCTGACACTCAGCATGAAATCTGGAGGCGCTGAATTTCAACCTCACGGAGATTC 377
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 598 CAGAAGCCCAATTCACGAGGGTTTTCAAGAGTTGTTGAGAACTTTGAATCAACCTGATT 657
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 378 CGGAGCTCAGATCCATGAAGGCTTCAGGAATCTCCCGTACCCTCAACCGCCAGACA 437
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 658 CTCATTTGCAATTAACACTACTGGTAACGGTTTATTTTGTCTGTAAGGTTTAAAAATTTGGTTG 717
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 438 GCCAGCTCCAGCTGACCCACCGCAATGCCTGTTCTCAGCGAGGGGCTGAAGCTAGTGG 497
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 718 ACAATTTCTTAGAGAGCTCAAGAACTATATCATAGTGAGGCTTTTACCGTTAAATTTG 777
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 498 ATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCAAGAGCGCTTCACTGTCAACTTCG 557
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 778 GTGATCTAGGAGAGCTAAAAGCAAAATTAATGATTTATGTAAGAGGAGGAGGAGGTA 837
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 558 GGGACCCGAGAGGCCAAGAAACAGATCAACGATTAAGTGGAGAGGGTACTCAAGGGA 617
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 838 AGATCGTTGACCTAGTTTAAAGAAATTAGATCGTGATACCGCTTTCGCGACTAGTTAACTATA 897
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 618 AATTTGTTGATTTGGTCAAGGAGCTTCACAGAGACACAGATTTTGTCTCTGTTGAATTAACA 677
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 898 TTTTTCACAGGGTAAGTGGGAAGCTCTTTCAGAGTTTAAAGTACTGAAGAGAGAGATTT 957
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 678 TCTTTTAAAGGCAATGGGAGAGACCTTTTGAAGTCAAGAGACACCGAGAGAGAGGACT 737
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 958 TTCATGTTGATCAAGTTTACTTCTCAAAGTTCCTCAATGATGAAAGAGCTGGGTATGTTCA 1017
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 738 TCCAGTGGACAGGTGACCCCGTGAAGTGCTATGATGAAGCGTTTAAAGCATGTTTA 797
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1018 ATATTCAACATTCGAAAAAATTAAGTTCTTTGGGTCTTTTAAATGAAGTATTTAGGTAACG 1077
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 798 ACATCCAGCACTGTGAAGAGCTGTCAGCTGGGTGCTGCTGATGAAATACCTGGCAATG 857
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1078 CTACTGCTATTTTCTTACCAGCAAGGTAAGCTTCAACATTTAGAGAAATGAGTTGA 1137
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 858 CCACCGGCATCTTCTTCTGCTGATGAGGGGAAACTTACAGACCTCGGAAATGAATCA 917
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1138 CTCATGACATTTACTTAAATTTTATAGAAACGAGGATTCGCTAGCGCTTCTCTGCACC 1197
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 918 CCCAGATATCATCACCAGTTCTTGAAATTAAGACAGAAAGTGTGCCAGCTTACATT 977
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1198 TGCCAAAGTTAAGTATCACCGGTACTTACGCTTAAATCTGTTTTAGGCCAGTTAGGTA 1257
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 978 TACCAAACTGTCCATTACTTGAACCTTATGATCTGAAGAGCGTCTCTGGGTCAACTGGCA 1037
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1258 TTACCAAGTTTCTTAAAGGTCGCCGATTTGAGTGGTGTACTGAAGAAGCTCCATTAA 1317
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1038 TCATTAAGGTTCTTCAAGCAATGGGGCTGACCTCTCCGGGGTCACAGAGGAGGACCCCTGA 1097
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1318 AATTGAGTAAAGCTGTTACAAAAGCCGCTTTAACTATTGATGAAAGGGGTACCGAGGCCG 1377

```


QY 358 CTCAAAAACCGACACAGTATCATCACCAAGACCAACATCGACTTTTAAATAAATTACTC 417
 Db 141 CCCAGACAGATATACATCCACCATGATCAGGATCACCAACCTTCAACAAGATCACCC 200
 QY 418 CAAATTTAGCCGAATTTGCTTTTCTTTGATAGACAAATAGCTCATCAAGTAATCTTA 477
 Db 201 CCAACCTGGCTGAGTTCGCCCTTCAGCCTATACCCAGCTGGCACACCACTCAACAGCA 260
 QY 478 CTAAATTTTATGCTGCTTTCTATTGCCACTGCTTTCGCCATGTTGAGTTAGGTA 537
 Db 261 CCAATATCTTCTTCCCCAGTGAGCTGCGTACAGGCTTTGCAATGCTTCCCTGGGGA 320
 QY 538 CTAAAGCCGATACCATGACGAGATTTTAGAAGGTTTAAAGTTTAAATTTGACCGAAATCC 597
 Db 321 CCAAGGCTGACACTCAGATGAATCTTGGAGGCTTGAATTTCAACCTCAGGAGATTC 380
 QY 598 CAGAACCCAAATTCACGAGGTTTCAAGAGTTGTTGAGAACTTTGAATCAACCTGATT 657
 Db 381 CGAGGCTCAGATCCATGAAGGCTTCCAGGAATCTCCCTGACCTCAACGAGCAGACA 440
 QY 658 CTCATTTGCAATTAACCTGCTGTTTATTTTCTGCTGAAGGTTTAAATTTGTTG 717
 Db 441 CCAAGCTCCAGCTGACCAACCGGCAATGCTTCTCAGGAGGCTGAGCTAGTGG 500
 QY 718 ACAATTTCTAGAACGCTCAAGAACTATATCATAGTGAAGCTTTTACCGTTAAATTTG 777
 Db 501 ATAGCTTTTGGAGGATGTTAAAGTTTGTACCACTCAGAGGCTTCACTGTCAACTTCG 560
 QY 778 GTGATCTGAGGAAGCTAAAAGCAAAATTAATGATTGTTGAGAAAGGACCCAGGGTA 837
 Db 561 GGGACACGAAGAGGCCAAGAACAGATCAACGATTACGTGAGAGGGTACTCAAGGGA 620
 QY 838 AGATCGTTGACCTAGTTAAGAAATAGATCGTATACCGCTTCCGACTAGTTAACTATA 897
 Db 621 AAATTTGTGGATTTGTCTAAGGAGCTTGACAGACACAGTTTTCCTGCTGTTGATTA 680
 QY 898 TTTTCTTTCAAGGCTAAGTGGGAACCTCTCTCGAGGTTTAAAGATCTGAAGAGGAAGT 957
 Db 681 TCTTCTTTAAAGGCAATGGGAGACCTTTTGAAGTCAAGGACACCCGAGGAAGGACT 740
 QY 958 TTTATGATGATCAAGTACTACTGTCAAAGTTCCAAATGATCAAGAACTGGTATGTTCA 1017
 Db 741 TCCAGCTGGACAGCTGACCACTGAAAGTGGCTTATGATCAAGGCTTTAGGCTATTTA 800
 QY 1018 ATATTCAACATTTGCAAAAATTAAGTTCTTGGCTCTTAAATGAAGTATTAGTAAAG 1077
 Db 801 ACATCCAGCACTGTAGAAGCTGCTCCAGCTGGGTCTGCTGATGAATACCTGGGCAATG 860
 QY 1078 CTACTGCTATTTTATTTTACACGAGAGGTAAGCTTCAACATTTAGAAATGAGTTGA 1137
 Db 861 CCACCGCCATCTTCTTCTGCTGATGAGGGGAACTACAGCACCTGGAAATGAACCTCA 920
 QY 1138 CTCATGACATTAATTAATAATTTTAGACAGAGGATCGTCTGAGCGCTTCTCTGCACC 1197
 Db 921 CCCAGCATATCATCAACAGTTCTTGGAAATGAAGACAGAGGCTGCCAGCTTACATTT 980
 QY 1198 TGCCAAAGTTAAGTATCACCGTACTTACGACTTAAATCTGTTTATAGCCAGTTAGGTA 1257
 Db 981 TACCCAACTGCTCCATTAAGTGAACCTATGATCTGAAGAGGCTCTGGTCAACTGGGCA 1040
 QY 1258 TTACCAAACTTTTCTTAACGCTGGGATTTGAGTGGTGTACTGAAGAGCTTCAATTAA 1317
 Db 1041 TCACTAAGTCTTTACGAAATGGGCTGACCTCTCCGGGCTCACAGAGGAGCCACCTGGA 1100
 QY 1318 AATTGATTAAGCTGTTCAAGAGGCTCTTAATCTATTGATGAAGGTTACCGAGGCGG 1377
 Db 1101 AGCTCTCCAAAGCCGTCGATAGGCTGTCTGCTGACCATGACGAGAGAGGGACTGAGCTG 1160
 QY 1378 CCGGCGCTATGTTCTTGGAGCTATTTCCTCAATGAGCATTCACAGAGGTTAAATTTAATA 1437
 Db 1161 CTGGGGCCATGTTTATAGAGGCCATACCATGCTATATCCCCCGAGGTTCAAGTTCAACA 1220

QY 1438 AACCATTCCTTTTCTGATGATCGAGACACATAAAGCCCATTTGTTTATGGCTAAGG 1497
 Db 1221 AACCCCTTTGCTCTTCTTAATGATTGAACAAATATACCAAGTCTCCCTCTCTCATGGGAAAG 1280
 QY 1498 TTGTCAACCAACTCAGAACTA 1519
 Db 1281 TGGTGAATCCCAACCAATA 1302
 RESULT 15
 LOCUS I07849 1434 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 2 from Patent EP 0103409.
 ACCESSION I07849
 VERSION I07849.1 GI:589437
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1434)
 AUTHORS Kawasaki, G.H. and Woodbury, R.G.
 TITLE Glycolytic promoters for regulated protein expression: protease inhibitor
 JOURNAL Patent: EP 0103409-A2 2 21-MAR-1984;
 FEATURES Location/Qualifiers
 source I. .1434
 BASE COUNT 369 a 432 c 293 g 0
 ORIGIN

Query Match 28.4%; Score 433.2; DB 6; Length 1434;
 Best Local Similarity 59.7%; Pred. No. 5.5e-83;
 Matches 729; Conservative 0; Mismatches 493; Indels 0; Gaps 0;
 QY 298 TGTGTGTAAAGTCTGTGTTTCCCACTCAAGGCCATGGAAGACCCTCAAGGGGAGCGCG 357
 Db 81 TGGCAGCGCTGTGCTGCTGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 140
 QY 358 CTCAAAAACCGACACAGTATCATCACCAAGACCAACATCGACTTTTAAATAAATTACTC 417
 Db 141 CCCAGACAGATATACATCCACCATGATCAGGATCACCAACCTTCAACAAGATCACCC 200
 QY 418 CAAATTTAGCCGAATTTGCTTTTCTTTGATAGACAAATAGCTCATCAAGTAATCTTA 477
 Db 201 CCAACCTGGCTGAGTTGCTGCTTCCAGCTATACCCAGCTGGCACACCACTCAACAGCA 260
 QY 478 CTAAATTTTATGCTGCTTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537
 Db 261 CCAATATCTTCTTCCCACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 320
 QY 538 CTAAAGCCGATACCATGACGAGATTTTAGAAGGTTTGAAGACTTTGAATCAACCTGATT 597
 Db 321 CCAAGGCTGACACTCAGATGAATCTTGGAGGCTTGAATTTCAACCTCAGGAGATTC 380
 QY 598 CAGAACCCAAATTCACGAGGTTTTCAGAGGTTGTTGAGAACTTTGAATCAACCTGATT 657
 Db 381 CGAGGCTCAGATCCATGAAGGCTTCCAGGAATCTCCCTGCTGCTGCTGCTGCTGCTGCT 440
 QY 658 CTCATTTGCAATTAACCTGCTGTTTAAAGTTTATTTTCTGCTGAGGTTTAAATTTGTTG 717
 Db 441 CCAAGCTCCAGCTGACCAACCGGCAATGCTTCTCAGCGAGGCTTCAAGCTAGTGG 500
 QY 718 ACAATTTCTAGAGAGCTCAAGAACTATATCATAGTGAAGCTTTTACCGTTAAATTTG 777
 Db 501 ATAGCTTTTGGAGGATGTTAAAGTTTGTACCACTCAGAGGCTTCACTGTCAACTTCG 560
 QY 778 GTGATCTGAGGAAGCTAAAAGCAAAATTAATGATTGTTGAGAAAGGACCCAGGGTA 837
 Db 561 GGGACACGAAGAGGCCAAGAACAGATCAACGATTACGTGAGAGGGTACTCAAGGGA 620
 QY 838 AGATCGTTGACCTAGTTAAGAAATAGATCGTATACCGCTTCCGACTAGTTAACTATA 897
 Db 621 AAATTTGTGGATTTGTCTAAGGAGCTTGACAGACACAGTTTTCCTGCTGTTGATTA 680

QY	898	TTTTTTTCAAGGGTAAGTGGGAAGCTCCTTTTCGAGGTTAAAGATPACTGAAGAGGAAGATT	957
Db	681	TCTTCTTTAAAGGCAAAATGGGAGAGACCTTTTGAAGTCAAGGACACCGAAGAGGACT	740
QY	958	TTTCATGTTGATCAAGGTACTACTGTCAAAGTTTCCAAGTATCAAAAGACTCGGTATGTTCA	1017
Db	741	TCCAGGTGGACCGTGAACCGTGAAAGTGCCTATGATGAAGCGTTTTAGGCATGTTTA	800
QY	1018	ATATTCAACATTTGCAAAAAATTAAGTCTTTGGGCTCTTATTAAATGAAGTATTTAGGTAAAG	1077
Db	801	ACATCCAGCACTGTAAAGAAGCTGCCAGCTGGGTGCTGATGAATAACCTGGGCAATG	860
QY	1078	CTACTGCTATTTTTTTTTTACCAGACNAGTAAGCTTCAACATTTAGAGAAATGAGTTGA	1137
Db	861	CCACCGCCATCTTCTCGTGCCTGATGAGGGGAAACTACAGCACTGGAAAAATGAATCA	920
QY	1138	CTCATGACATTTACTAAATTTTTAGAGAACGAGGATCGTGTAGGCGTCTCTCGCAAC	1197
Db	921	CCCACGATATCATCACCAGTTCTGGAAATGAAGACAGAAAGTCTGCCACGCTTACATT	980
QY	1198	TGCCAAAGTTAAGTATCACCGGTACTTACGACTTTAAAATCTGTTTTAGGCCAGTTAGGTA	1257
Db	981	TACCCAAACTGTCCATTTACTGGAACCTATGATCTGAAGAGCGTCTGGGTCAACTGGCA	1040
QY	1258	TTACCAAAAGTTTTTTTCTAACCGTCCCATTTAGTGTGTGTACTGAAGAAGTCCATTAA	1317
Db	1041	TCACTAAGGCTTTTCAGCAATGGGGCTGACCTCTCCGGGGTCAAGAGGAGGCCACCTGA	1100
QY	1318	AATTGAGTAAAGCTGTTCCAAAAAGCCGTCTTTAACTATTGATGAAAAGGTTACCGAGGCCG	1377
Db	1101	AGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATCGACGAGAAAGGACTGAAGCTG	1160
QY	1378	CCGGCGCTATGTTCTGGAAGCTATTCCATGAGCAATTCACCAGAAAGTTTAAATTAAATA	1437
Db	1161	CTGGGGCCCATGTTTTTAGAGGCCATACCATGTCTATCCCCCGGAGGTCAAGTTCAACA	1220
QY	1438	AACCATTCGTTTTTCTGATGATCCAGCAGAACACTAAAGCCCATTTGTTATGGGTAAAGG	1497
Db	1221	AACCTTTGTCTCTTATGATGATGAACAAATACCAAGTCTCCCTCTTCATGGGAAAAG	1280
QY	1498	TTGTCAACCCCAACTCAGAAGTA	1519
Db	1281	TGGTGAATCCCACCCCAAAAAATA	1302

Search completed: December 6, 2002, 22:41:32
Job time : 2703 secs

